DOCUMENT RESUME

ED 141 834 CS 203 539

AUTHOR Kates, Jack
TITLE The Santa Barbara County Experiment on Individualizing Ninth Grade English Composition: The Kates Method.
INSTITUTION Santa Barbara County Schools, Calif.; Writing Teachers In-Service Program, Cypress, Calif.
PUB DATE 77
NOTE 76p.

EDRS PRICE MF-$0.83 HC-$4.67 Plus Postage.
DESCRIPTORS Composition (Literary); *Composition Skills (Literary); Educational Research; *English Instruction; *Expository Writing; Grade 9; Junior High Schools; *Program Evaluation; Sequential Approach; *Teaching Techniques
IDENTIFIERS *Teacher Student Conferences

ABSTRACT

The effectiveness of the Kates method of teaching composition, which involves regular, sequential writing assignments and teacher/student conferences, was investigated using an experimental group of 142 ninth-grade students in six classes and a control group of 78 students in two classes. The experimental classes wrote approximately 14 essays in class during the course of the semester and engaged in five-minute conferences on alternate weeks. The control group was less structured, completing 10 to 12 paragraph-writing assignments, in and out of class; they received grades for only four or five assignments and had no conferences. An essay pretest and an essay posttest were administered to all students and were scored by an outside reader for content, organization, mechanics, and sentence structure. Analyses of results indicated gains in most areas tested, but only for the experimental classes. (Author/AA)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED141834

THE SANTA BARBARA COUNTY EXPERIMENT

ON INDIVIDUALIZING NINTH GRADE ENGLISH COMPOSITION:

THE KATES METHOD

Funded by

Santa Barbara County Schools
Office of the Superintendent
Paul Jillson, Director of
Curriculum and Instructional Services

and

Writing Teachers In-Service Program
6505 Jaluit Street, Cypress, California 90630
Jack Kates, Director

by Jack Kates

CS 203 539

TABLE OF CONTENTS

		Page
LIST OF TABLES		iii
ACKNOWLEDGMENTS		iv
I.	INTRODUCTION	1
II.	PROBLEM	3
III.	SIGNIFICANCE	4
IV.	HYPOTHESIS	4
V.	METHOD	6
VI.	RESULTS	12
VII.	DISCUSSION OF RESULTS	32
VIII.	CONCLUSION	39
IX.	SUMMARY	47
X.	RECOMMENDATIONS	49

LIST OF TABLES

Table		Page
1	Pre-Post Scores - Content - Experimental Group	12
2	Pre-Post Scores - Organization - Experimental Group. . . .	13
3	Pre-Post Scores - Mechanics - Experimental Group	14
4	Pre-Post Scores - Sentence Structure - Experimental Group.	15
5	Pre-Post Scores - Composite (Final) - Experimental Group .	16
6	Pre-Post Scores - Content - Control Group	17
7	Pre-Post Scores - Organization - Control Group	18
8	Pre-Post Scores - Mechanics - Control Group	19
9	Pre-Post Scores - Sentence Structure - Control Group . . .	20
10	Pre-Post Scores - Composite (Final) - Control Group . . .	21
11	Pre-Post Individual Scores - Content - Experimental Group.	22
12	Pre-Post Individual Scores - Organization - Experimental Group	23
13	Pre-Post Individual Scores - Mechanics - Experimental Group	24
14	Pre-Post Individual Scores - Sentence Structure - Experimental Group	25
15	Pre-Post Individual Scores - Composite (Final) - Experimental Group	26
16	Pre-Post Individual Scores - Content - Control Group . . .	27
17	Pre-Post Individual Scores - Organization - Control Group.	28
18	Pre-Post Individual Scores - Mechanics - Control Group . .	29
19	Pre-Post Individual Scores - Sentence Structure - Control Group	30
20	Pre-Post Individual Scores - Composite (Final) - Control Group	31
21	Pre-Post Test Achievement Chart: Composite (Final) Group Scores	31A

ACKNOWLEDGMENTS

First of all, I would like to thank all of the ninth grade instructors and their students who participated in this experiment, at Lompoc Senior High School and at the control school. Secondly, I would like to thank Mrs. Sue Sciame, for grading all of the papers most conscientiously, and Richard L. Mercer, of Golden West College, for computerizing the results. In addition, I would like to express my gratitude to Mr. Paul Jillson, director of curriculum and instructional services at Santa Barbara County Schools, for providing part of the funding and for furnishing the control classes. Finally--and most importantly--I would like to express my sincerest gratitude to Mr. Gary Ingraham, formerly the English department chairman at Lompoc Senior High School and currently the vice principal at Oakdale Joint Union High School, not only for initiating, implementing, and overseeing the project, but also for obtaining part of the funding from the Santa Barbara County Schools.

THE SANTA BARBARA COUNTY EXPERIMENT ON INDIVIDUALIZING NINTH GRADE ENGLISH COMPOSITION: THE KATES METHOD

I. INTRODUCTION

The report on this experiment dealing with individualized instruction in the teaching of English composition follows the researcher's previous work in this area of concentration, work that includes not only published and unpublished research but also international teacher-training seminars for English instructors and language arts specialists.

Previous research by this writer reveals that the individualized conference method of teaching composition is superior to the traditional method used by most instructors, the traditional method being defined as the returning of student essays with symbols and written comments by the instructor as a means of feedback instead of individual teacher-student conferences. Although the subjects of his previous research were college freshmen, the researcher assumed that the same teaching-learning principles could be applied to students of any age or grade level, including junior or senior high school students. The assumption formed the basis of his method of teaching writing skills, a method--now known as the "Kates Method"--which he has shared with thousands of English instructors and language arts specialists who have attended his international in-service seminars, offered annually throughout North America.

Even though the assumption appeared valid enough, and even though hundreds of teachers reported successful results with the Kates Method shortly after trying it during the first year of his seminar, the assumption and the method were soon challenged. The challenge arose from the

English Department Chairman of Lompoc Senior High School because he felt that at the high school level, certain factors--particularly classroom management and control--might actually hinder the instructor during his efforts to provide individual conferences for all the students. These factors were negligible or non-existent at the college level, where the researcher had done his previous work. The department chairman, also being research oriented, felt that the consultant's assumption could not be considered valid unless it were actually tested on the sample specified, that is, on high school students. The department chairman, therefore, asked the consultant, who had presented the in-service seminar for the school district, if he would care to design an experiment that would test the reliability of the Kates Method. The subjects of the experiment would be entering ninth-graders during the fall semester of 1975.

Realizing that he did need new data to support or to disclaim his previous research, the consultant agreed not only to assist in the design of the experiment but also to arrange for the computer services and to finance part of the project through Writing Teachers In-Service Program. The remainder of the financing would be arranged by the department head, through the curriculum department of Santa Barbara County Schools.

This report, then, describes the experiment on the ninth-graders involved in the project: six experimental classes at Lompoc Senior High School and two control classes at a neighboring school in a different school district but within the same county.

II. PROBLEM

The problem was to test the Kates Method of teaching English composition on several classes of entering ninth grade students, for an entire semester. The Kates Method--an individualized conference technique--was an outgrowth of the researcher's previous work on individualization, and was being accepted by hundreds of English instructors who had attended special teacher-training seminars offered by Writing Teachers In-Service Program. Although the method had been tested thoroughly at the college level, where classroom control problems were virtually non-existent, it had never been tested at the junior or senior high school level, where control problems were sometimes so severe that the teacher found it necessary to maintain a constant surveillance over his class, thus precluding any form of individual instruction whatsoever. The problem, therefore, was to test the Kates Method of individualization on a large sample of high school students.

III. SIGNIFICANCE

If such a study could be conducted on a large sample under controlled conditions, there might then be generally available data on the "actual" writing skills of high school students. This data might serve either to validate or to repudiate the reliability of the Kates Method of teaching English composition, a method that had already been accepted by hundreds of junior and senior high school teachers, and one that might be accepted in the future by thousands more.

IV. HYPOTHESIS

Because all classes tested were considered to be fairly heterogeneous in makeup*, it was hypothesized, first of all, that a significant range of abilities would be evident among the students in each class as well as among the various classes, both in the pre-and-post test. This range would reflect the various skills tested, skills actually demonstrated by the students in their essay writing. The specific skills to be tested included content, organization, mechanics, and sentence structure. The final composite score was an average of the four skills tested.

*(The groups were considered to be heterogeneous because no attempt had been made to group them according to ability level except for the initial screening at Lompoc High School, when the lowest seven percent were separated and placed in a special remedial class.)

Secondly, it was hypothesized that the post-test would reveal a significant achievement or gain in most areas tested, both for the experimental group and the control. The gain, like the other scores, would be calculated on the stanine and confirmed by a 2 tail test using the z-ratio.

Thirdly, it was hypothesized that because of the Kates Method of individualized instruction, the experimental classes would achieve, in

all areas tested, greater and more significant gains than those achieved by the control classes.

Finally, it was hypothesized that the overall mean achievement of the experimental group (the experimental classes as a whole) based on the "composite final score," would be comparable to that exhibited by the entering college freshmen who had been tested three years earlier, during the Kates Survey of College Freshman Composition Writing Skills: Sixteen First-Semester English Composition Classes from Eleven Community Colleges and Universities, ERIC, ED 077 498. The mean gain of the 285 freshmen tested had been 1.804 stanines at a significance level of 0%.

V. METHOD

The design of the experiment was established to compare, through the use of mean stanine scores as well as the 2 tail test involving the z-ratio, the actual gains in writing skills and ratios of gains among eight comparable, fairly heterogeneous groups of entering ninth grade students during a semester's course in English composition. The experimental groups were to include six classes at Lompoc Senior High School, all furnished by its English department chairman. The control groups were to include two classes at a neighboring school in the same county, furnished by the county curriculum director. Both the pre-and-post test were to be based on a choice of two questions, and both tests were to be administered in class under controlled conditions. The pretest was to be administered to the untreated classes, during the first week of school; the post-test would be administered at the end of the course, eighteen weeks later.

After the pretests had been administered, they were to be collected--ungraded--coded by class and school, and then locked in a vault. At the end of the course, the post-test, identical to the pre, would be administered, after which the papers would again be collected and coded. Because logistics precluded all of the experimental papers being graded (over 500), sizeable samples to represent the six experimental classes would then be drawn from both tests, pre-and-post. However, the papers from one entire experimental class would also be included in the final sample. Finally, all of the papers to be graded, from all of the classes, both pre-and-post tests, would be mixed together, so that the reader would not know whose paper she was grading nor from what class it was taken. To ensure anonymity, the sheet containing the instructions covered the grading sheet containing the student's identification.

The papers to be graded included only those written by students who had taken both the pre-and-post tests. The resultant scores would be used to compute the 2 tail z-ratio, which is more reputable than the 1 tail test. (By using the 2 tail test, the researcher avoids the accusation that the low students were dropped in order to show improvement. If the results are significant with a 2 tail test, they are always significant with a 1 tail test.)

The population samples, both experimental and control, were fairly comparable to each other in socio-economic makeup and very comparable in size. The experimental classes represented a cross section of the country as a whole, for they included students from low, average, and high average income families, as well as students of all colors and backgrounds. Students from the control classes, however, generally came from a more affluent, predominately white, rural background. The size of all classes ranged from 34 to 48, and every class was team-taught by two instructors.

The major difference between the experimental and control group would be the method of instruction. The experimental group would use the Kates Method, focusing on the individual conference, whereas the control group would use the traditional method, in which the teacher marked the papers and wrote comments on them but did not supply individual conferences. The Kates Method, in addition, stressed weekly, in-class writing assignments for all students. The traditional method left the amount of writing up to the individual instructor.

Although the instructors at Lompoc Senior High School had been introduced to the Kates Method in May, just before the end of the previous school year, the department chairman thought it necessary that his ninth grade teachers receive additional, intensive instruction in the conference

method, so that they would become most effective in it. Accordingly, the consultant agreed to demonstrate the method again, for three days at the beginning of the fall semester, during the regular school hours. In addition, he agreed to offer suggestions on the sequencing of writing assignments.

As the first essay of the new course, the pretest was administered to all students during the first week of school, after which all papers were collected, coded, and stored in a vault for safekeeping. Following the pretest, the students in the experimental classes wrote a second essay in class, a process analysis theme, on which they would be "conferenced" by the consultant. Immediately afterward, for three consecutive days, the consultant provided almost every student in the experimental classes with an individual conference on the process theme. After the teachers had observed the consultant for two days, they, themselves, practiced the method on their own students under his observation. While each student was being conferenced on his paper, the rest of the class was engaged in writing still another essay, its third in two weeks.

The Kates Method suggested that each student write one essay per week, with three days being devoted to the writing. On the first day, the student received specific instructions on the type of writing to be done and on what was expected. He then proceeded to write the outline, followed by a first draft. On the second day, he was to rewrite his paper, carefully, for the second draft. After completing the second draft, he was to work in an "editing" group, proofreading his peers' papers, for mechanical errors, while they in turn proofread his. Finally, on the third day, he was to write the final draft and to turn it in to his instructor. The final draft was the one that his instructor would use for the conference, usually on the following week. If some of the students

finished their assignments ahead of schedule, they would have writing-related activities, which included exercises on grammar, punctuation, usage, or whatever the instructor felt was necessary to improve writing skills.

During the semester's course, each student in the experimental group wrote approximately fourteen essays, most of them varying from one to five paragraphs in length. Themes included process analysis, at the beginning of the course, and then persuasion through thesis-and-supporting evidence. Each student received from his instructor one five-minute conference on every other paper, usually on alternate weeks. The conference covered the four items graded--content, organization, mechanics (spelling and punctuation), and sentence structure (including diction). However, the focus was on organization because the department chairman and his faculty felt it to be the most important item. Although the conference was supposed to last approximately five minutes and was supposed to be given only on the final draft, the instructors in two experimental classes, IV and V, felt it was necessary to provide slightly longer conferences, not on the final draft but on the first one. Afterward, they provided brief conferences on the final draft, along with the grade. Their students were the only ones who received conferences on the first draft, and thus feedback on both. The idea of three drafts, however, was ultimately discarded by all the experimental instructors because they felt that their students complained too much over writing the third draft; and the logistics of the "editing" groups had not been implemented. Only two drafts, then, were required. All writing, however, was done in class.

By contrast, the control groups were much less structured in their writing. There was no emphasis on the "final" draft, and although the students were supposed to do about ten or twelve assignments from a

programmed textbook, many of them were done out of class, and only four or five were actually graded. Grading consisted of symbols and written comments on the papers. No individual conferences were provided unless they were requested by the student.

One week prior to the end of the course, the post-test, identical to the pre-test, was administered to all students. Again, rigidly controlled conditions prevailed. The questions were unknown to the students ahead of time. The time limit, again, was fifty minutes. Under the supervision of the department head, the post-test papers were all collected and coded; then samples representing the experimental classes were drawn from both the pre-and-post tests. Finally, all the papers to be graded--both pre-and-post tests, both experimental and control--were mixed together for the reader to grade. The total sample numbered 440 papers, 284 from the experimental classes and 156 from the control.

In order to ensure objectivity in the grading and to prevent bias, a disinterested party was engaged as the outside reader who would do all of the grading. An added factor of standardization was thereby built in. Moreover, to ensure as much accuracy as possible on the Kates Grading Scale, the reader received three days of intensive instruction from the consultant in the use of that scale. Together they read and graded over 50 papers, eleven percent of the total sample, until there was no discrepancy between the consultant's grades and those assigned by the reader on any given paper. Standardization in the grading was thereby ensured. Finally, to prevent fatigue from affecting the accuracy of the grading, the reader graded no more than ten papers per day.

Grading criteria on the Kates Grading Scale consist of four equally-weighted items: content, organization, mechanics, and sentence structure.

Content involves what is said, not only in its substance, but also in its significance, as well as in its factual support. Generalizations must be supported by logical, specific evidence, or they are invalid. Organization includes such requirements as a brief introductory paragraph containing an explicit thesis statement referring to the topic or essay question, coherence between sentences and between paragraphs, development of individual paragraphs--when needed--through concrete detail, and a brief concluding paragraph which unifies the essay as a whole. Mechanics include spelling and punctuation. Finally, sentence structure includes correct syntax or word order, appropriate diction, a variation of sentence patterns, and correct grammar and usage.

Each equally-weighted criterion is graded on the stanine (standard scale of nine), with 1 representing the lowest possible grade, 2 and 3 representing below average work; 4, 5, and 6 representing varying degrees of the norm (C-, C, C+); and 7, 8, and 9 equivalent to the above average grades of B- to B, B+ to A-, and A respectively. The grades are then averaged by adding them and dividing by four. The resultant composite stanine score--final score--can be used for comparisons in many ways, including both individual and group achievements.

After all of the papers had been graded, they were collected by the department head, decoded, separated by test (pre-or-post), by group, and by class, and then arranged alphabetically by students' names, in order to record the scores. The scores included not only the composite (final) score, but also the score for each item tested: content, organization, mechanics, and sentence structure. There was thus available a profile on writing skills for each student, for each class, and for each group.

Finally, the papers, along with the records, were sent to the consultant, who then computerized the results.

VI. RESULTS

EXPERIMENTAL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT: CONTENT - STANINE SCORES

SCHOOL	SAMPLE SIZE	MAX.	MIN.	RANGE	MEAN	VARIANCE	STANDARD DEVIATION	MEAN DEVIATION	MEDIAN	MODE	STANINE GAIN	z-RATIO	SIGNIFICANT AT
LOMPOC-I	pre- 18	7	1	6	4.44	2.03	1.42	1.12	5	5			
	post-	9	2	7	7.33	3.65	1.91	1.44	8	9	2.89	0.030	5%
LOMPOC-II	pre- 43	8	1	7	4.16	2.62	1.62	1.31	4	4			
	post-	9	1	8	7.09	4.94	2.22	1.75	8	9	2.93	0.012	5%
LOMPOC-III	pre- 19	7	2	5	4.58	2.92	1.71	1.50	5	3			
	post-	9	4	5	6.00	2.33	1.53	1.16	6	5	1.42	0.016	5%
LOMPOC-IV	pre- 16	6	1	5	3.13	1.98	1.41	1.16	3	2			
	post-	9	2	7	5.81	5.10	2.26	1.81	6	5	2.69	0.034	5%
LOMPOC-V	pre- 25	7	1	6	3.52	2.43	1.56	1.30	3	3			
	post-	9	2	7	5.96	4.46	2.11	1.72	6	7	2.44	0.021	5%
LOMPOC-VI	pre- 21	8	1	7	4.05	3.55	1.88	1.49	3	3			
	post-	9	1	8	4.76	4.89	2.21	1.72	5	5	0.71	0.260	NOT
TOTAL	pre- 142	8	1	7	4.01	2.73	1.65	1.36	4	3			
	post-	9	1	8	6.29	4.94	2.22	1.87	7	9	2.28	0.000	0%

Table 1

EXPERIMENTAL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT: ORGANIZATION - STANINE SCORES

SCHOOL	SAMPLE SIZE	MAX.	MIN.	RANGE	MEAN	VARIANCE	STANDARD DEVIATION	MEAN DEVIATION	MEDIAN	MODE	STANINE GAIN	z-RATIO	SIGNIFICANT AT
LOMPOC-I	pre- 18	6	1	5	3.72	1.74	1.32	1.01	4	4			
	post-	9	2	7	7.28	3.62	1.90	1.36	8	7, 8, 9	3.56	0.030	5%
LOMPOC-II	pre- 43	7	1	6	3.91	2.32	1.52	1.18	4	4			
	post-	9	2	7	7.14	3.88	1.97	1.51	8	8	3.23	0.012	5%
LOMPOC-III	pre- 19	8	2	6	4.26	3.20	1.79	1.41	4	3			
	post-	9	4	5	6.11	2.32	1.52	1.29	6	5	1.84	0.026	5%
LOMPOC-IV	pre- 16	6	1	5	3.25	1.80	1.34	1.06	3	3			
	post-	9	2	7	6.00	3.87	1.97	1.50	6	6	2.75	0.034	5%
LOMPOC-V	pre- 25	9	1	8	3.64	3.32	1.82	1.36	3	3			
	post-	9	2	7	5.84	4.47	2.12	1.71	6	7	2.20	0.021	5%
LOMPOC-VI	pre- 21	7	1	6	3.48	1.76	1.33	0.97	3	3			
	post-	9	2	7	5.05	3.75	1.94	1.58	5	7	1.57	0.021	5%
TOTAL	pre- 142	9	1	8	3.75	2.40	1.55	1.22	3	3			
	post-	9	2	7	6.35	4.22	2.05	1.73	7	7	2.61	0.000	0%

Table 2

EXPERIMENTAL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT: MECHANICS - STANINE SCORES

SCHOOL	SAMPLE SIZE	MAX.	MIN.	RANGE	MEAN	VARIANCE	STANDARD DEVIATION	MEAN DEVIATION	MEDIAN	MODE	STANINE GAIN	z-RATIO	SIGNIFICANT AT
LOMPOC-I	pre- 18	9	1	8	3.61	10.25	3.20	2.97	2	1			
	post-	9	1	8	4.06	9.35	3.06	2.62	3	1	0.44	0.672	NOT
LOMPOC-II	pre- 43	9	1	8	2.79	5.55	2.36	1.91	2	1			
	post-	8	1	7	3.56	7.92	2.81	2.57	2	1	0.77	0.169	NOT
LOMPOC-III	pre- 19	7	1	6	2.42	3.81	1.95	1.58	1	1			
	post-	8	1	7	3.47	6.37	2.52	2.18	3	1	1.05	0.147	NOT
LOMPOC-IV	pre- 16	8	1	7	2.06	3.93	1.98	1.35	1	1			
	post-	9	1	8	4.13	5.72	2.39	1.89	4	1, 3, 5	2.06	0.018	5%
LOMPOC-V	pre- 25	8	1	7	2.28	3.21	1.79	1.40	1	1			
	post-	9	1	8	3.92	9.16	3.03	2.79	3	1	1.64	0.006	1%
LOMPOC-VI	pre- 21	8	1	7	2.19	4.66	2.16	1.70	1	1			
	post-	9	1	8	3.33	8.63	2.94	2.41	2	1	1.14	0.148	NOT
TOTAL	pre- 142	9	1	8	2.58	5.22	2.29	1.85	1	1			
	post-	9	1	8	3.70	7.77	2.79	2.48	3	1	1.12	0.000	0%

Table 3

EXPERIMENTAL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT: SENTENCE STRUCTURE - STANINE SCORES

SCHOOL	SAMPLE SIZE	MAX.	MIN.	RANGE	MEAN	VARIANCE	STANDARD DEVIATION	MEAN DEVIATION	MEDIAN	MODE	STANINE GAIN	z-RATIO	SIGNIFICANT AT
LOMPOC-I	pre- 18	9	1	8	5.50	8.38	2.90	2.61	7	7			
	post-	9	2	7	6.56	4.97	2.23	1.88	8	8	1.06	0.219	NOT
LOMPOC-II	pre- 43	9	1	8	4.63	6.86	2.62	2.30	4	2			
	post-	9	1	8	6.91	3.80	1.95	1.48	7	8	2.28	0.012	5%
LOMPOC-III	pre- 19	9	1	8	4.79	6.18	2.49	2.14	5	3, 6			
	post-	9	3	6	6.89	2.32	1.52	1.20	7	8	2.11	0.024	5%
LOMPOC-IV	pre- 16	9	1	8	3.81	6.83	2.61	2.16	3	2			
	post-	9	2	7	7.38	3.72	1.93	1.36	8	8	3.56	0.034	5%
LOMPOC-V	pre- 25	9	1	8	3.84	7.89	2.81	2.38	3	1			
	post-	9	1	8	6.64	5.91	2.43	1.95	8	8	2.80	0.021	5%
LOMPOC-VI	pre- 21	9	1	8	3.38	7.75	2.78	2.32	2	1			
	post-	9	1	8	5.86	6.23	2.50	2.05	6	5, 9	2.48	0.020	5%
TOTAL	pre- 142	9	1	8	4.35	7.45	2.73	2.45	3	1			
	post-	9	1	8	6.71	4.49	2.12	1.72	7	8	2.37	0.000	0%

EXPERIMENTAL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT: COMPOSITE (FINAL) - STANINE SCORES

SCHOOL	SAMPLE SIZE	MAX.	MIN.	RANGE	MEAN	VARIANCE	STANDARD DEVIATION	MEAN DEVIATION	MEDIAN	MODE	STANINE GAIN	z-RATIO	SIGNIFICANT AT
LOMPOC-I	pre- 18	7	1	6	4.17	3.09	1.76	1.41	4	4, 5			
	post-	9	2	7	6.28	3.04	1.74	1.31	6	6	2.11	0.029	5%
LOMPOC-II	pre- 43	8	1	7	3.74	3.39	1.84	1.54	3	2			
	post-	9	1	8	6.09	3.32	1.82	1.36	6	6, 7	2.35	0.012	5%
LOMPOC-III	pre- 19	7	1	6	3.95	3.16	1.78	1.53	4	2, 5			
	post-	8	3	5	5.53	2.15	1.47	1.24	5	5	1.58	0.022	5%
LOMPOC-IV	pre- 16	7	1	6	2.94	2.86	1.69	1.30	2	2			
	post-	8	2	6	5.63	2.65	1.63	1.22	6	6	2.69	0.034	5%
LOMPOC-V	pre- 25	6	1	5	3.20	2.33	1.53	1.34	3	2			
	post-	9	1	8	5.40	5.00	2.24	1.86	5	4	2.20	0.021	5%
LOMPOC-VI	pre- 21	7	1	6	3.19	2.96	1.72	1.41	2	2			
	post-	8	1	7	4.62	4.35	2.09	1.70	4	4	1.43	0.002	1%
TOTAL	pre- 142	8	1	7	3.56	3.07	1.75	1.52	3	2			
	post-	9	1	8	5.65	3.66	1.91	1.56	6	6, 7	2.09	0.000	0%

Table 5

CONTROL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT: CONTENT - STANINE SCORES

SCHOOL	SAMPLE SIZE	MAX.	MIN.	RANGE	MEAN	VARIANCE	STANDARD DEVIATION	MEAN DEVIATION	MEDIAN	MODE	STANINE GAIN	z-RATIO	SIGNIFICANT AT
GROUP I	Pre- 44	8	2	6	5.64	1.63	1.28	1.02	6	5			
	Post-	9	2	7	5.23	1.90	1.38	1.09	5	4	-0.41	0.147	NOT
GROUP II	Pre- 34	9	3	6	6.06	2.91	1.70	1.42	6	4			
	Post-	9	3	6	5.32	2.41	1.55	1.24	5	4	-0.74	0.057	10%
TOTAL	Pre- 78	9	2	7	5.82	2.20	1.48	1.20	6	5			
	Post-	9	2	7	5.27	2.10	1.45	1.15	5	4	-0.55	0.020	5%

Table 6

CONTROL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT: ORGANIZATION - STANINE SCORES

SCHOOL	SAMPLE SIZE	MAX.	MIN.	RANGE	MEAN	VARIANCE	STANDARD DEVIATION	MEAN DEVIATION	MEDIAN	MODE	STANINE GAIN	z-RATIO	SIGNIFICANT AT
	Pre-	8	2	6	5.34	2.42	1.55	1.27	5	6			
GROUP I	44												
	Post-	8	3	5	4.64	1.77	1.33	1.05	5	5	-0.70	0.015	5%
	Pre-	9	3	6	5.62	3.70	1.92	1.61	5	5			
GROUP II	34												
	Post-	9	2	7	5.32	3.44	1.85	1.54	5	4	-0.29	0.522	NOT
	Pre-	9	2	7	5.46	2.95	1.72	1.41	5	5			
TOTAL	78												
	Post-	9	2	7	4.94	2.58	1.61	1.22	5	4	-0.53	0.050	5%

Table 7

CONTROL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT: MECHANICS - STANINE SCORES

SCHOOL	SAMPLE SIZE	MAX.	MIN.	RANGE	MEAN	VARIANCE	STANDARD DEVIATION	MEAN DEVIATION	MEDIAN	MODE	STANINE GAIN	z-RATIO	SIGNIFICANCE AT
	Pre-	9	1	8	4.93	7.93	2.82	2.44	5	1			
GROUP I	44												
	Post-	9	1	8	4.43	7.65	2.77	2.50	5	1	-0.50	0.401	NOT
	Pre-	9	1	8	5.03	6.82	2.61	2.21	6	7			
GROUP II	34												
	Post-	9	1	8	3.88	7.20	2.68	2.39	3	2	-1.15	0.069	10%
	Pre-	9	1	8	4.97	7.35	2.91	2.34	5	1			
TOTAL	78												
	Post-	9	1	8	4.19	7.43	2.73	2.49	3	1	-0.78	0.074	10%

Table 8

CONTROL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT: SENTENCE STRUCTURE - STANINE SCORES

SCHOOL	SAMPLE SIZE	MAX.	MIN.	RANGE	MEAN	VARIANCE	STANDARD DEVIATION	MEAN DEVIATION	MEDIAN	MODE	STANINE GAIN	z-RATIO	SIGNIFICANCE AT
	Pre-	9	1	8	6.09	6.32	2.51	1.90	7	6, 8			
GROUP I	44												
	Post-	9	2	7	6.36	3.63	1.91	1.59	7	7	0.27	0.567	NOT
	Pre-	9	1	8	7.29	3.12	1.77	1.24	8	8			
GROUP II	34												
	Post-	9	2	7	6.09	3.54	1.88	1.57	6	8	-1.21	0.006	1%
	Pre-	9	1	8	6.62	5.23	2.29	1.72	7	8			
TOTAL	78												
	Post-	9	2	7	6.24	3.56	1.89	1.60	7	7, 8	-0.37	0.271	NOT

Table 9

CONTROL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT: COMPOSITE (FINAL) - STANINE SCORES

SCHOOL	SAMPLE SIZE	MAX.	MIN.	RANGE	MEAN	VARIANCE	STANDARD DEVIATION	MEAN DEVIATION	MEDIAN	MODE	STANINE GAIN	z-RATIO	SIGNIFICANT AT
	Pre-	8	1	7	5.30	2.96	1.72	1.37	6	6			
GROUP I	44												
	Post-	8	2	6	5.02	1.98	1.41	1.16	5	4	-0.27	0.416	NOT
	Pre-	9	2	7	5.94	2.78	1.67	1.31	6	7			
GROUP II	34												
	Post-	9	3	6	5.15	2.80	1.67	1.46	5	4	-0.79	0.043	5%
	Pre-	9	1	8	5.58	2.95	1.72	1.39	6	6			
TOTAL	78												
	Post-	9	2	7	5.08	2.31	1.52	1.29	5	4	-0.50	0.056	NOT

Table 10

EXPERIMENTAL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT--INDIVIDUAL STANINE SCORES-- CONTENT

SCHOOL	SAMPLE SIZE	MEAN SCORE	STANINE GAIN	z-RATIO	Below Average			Average			Above Average		
					1	2	3	4	5	6	7	8	9
LOMPOC-I	pre-18	4.44			1	0	4	2	10	2	1	0	0
	post-	7.33	2.89	0.030	0	2	0	0	2	1	5	2	7
LOMPOC-II	pre-43	4.16			1	6	9	10	7	7	2	1	0
	post-	7.09	2.93	0.012	1	2	1	2	4	2	6	10	15
LOMPOC-III	pre-19	4.58			0	2	5	2	3	4	3	0	0
	post-	6.00	2.93	0.012	0	0	0	2	7	5	1	2	2
LOMPOC-IV	pre-16	3.13			1	6	3	3	2	1	0	0	0
	post-	5.81	2.69	0.034	0	1	2	1	4	3	0	2	3
LOMPOC-V	pre-25	3.52			2	5	7	3	6	1	1	0	0
	post-	5.96	2.44	0.021	0	2	2	1	6	2	7	1	4
LOMPOC-VI	pre-21	4.05			1	1	10	2	3	1	1	2	0
	post-	4.76	0.71	0.260	3	0	3	2	5	4	2	1	1
TOTAL	pre-142	4.01			6	20	38	22	31	16	8	3	0
	post-	6.29	2.28	0.000	4	7	8	8	28	17	21	18	32

Table 11

EXPERIMENTAL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT--INDIVIDUAL STANINE SCORES-- ORGANIZATION

SCHOOL	SAMPLE SIZE	MEAN SCORE	STANINE GAIN	z-RATIO	Below Average 1	2	3	4	Average 5	6	Above Average 7	8	9
LOMPOC-I	pre- 18	3.72			1	2	4	7	2	2	0	0	0
	post-	7.28	3.56	0.030	0	1	0	1	1	0	5	5	5
LOMPOC-II	pre- 43	3.91			2	6	9	11	8	3	3	0	0
	post-	7.14	3.23	0.012	0	2	0	6	0	2	8	14	11
LOMPOC-III	pre- 19	4.26			0	2	7	3	2	2	2	1	0
	post-	6.11	1.84	0.026	0	0	0	2	7	3	2	4	1
LOMPOC-IV	pre- 16	3.25			1	4	5	3	2	1	0	0	0
	post-	6.00	2.75	0.034	0	1	1	2	1	4	3	3	1
LOMPOC-V	pre- 25	3.64			2	4	8	5	3	1	1	0	1
	post-	5.84	2.20	0.021	0	2	2	3	3	4	7	0	4
LOMPOC-VI	pre- 21	3.48			1	1	13	2	2	1	1	0	0
	post-	5.05	1.57	0.021	0	2	4	2	4	3	5	0	1
TOTAL	pre- 142	3.75			7	19	46	31	19	10	7	1	1
	post-	6.35	2.61	0.000	0	8	7	16	16	16	30	26	23

Table 12

EXPERIMENTAL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT--INDIVIDUAL STANINE SCORES-- MECHANICS

SCHOOL	SAMPLE SIZE	MEAN SCORE	STANINE GAIN	z-RATIO	Below Average			Average			Above Average		
					1	2	3	4	5	6	7	8	9
LOMPOC-I	Pre- 18	3.61			9	2	0	0	0	1	3	2	1
	Post-	4.06	0.44	0.672	5	3	2	1	1	2	0	1	3
LOMPOC-II	Pre- 43	2.79			19	9	3	2	2	4	2	0	2
	Post-	3.56	0.77	0.169	21	1	0	4	4	3	4	6	0
LOMPOC-III	Pre- 19	2.42			10	2	3	0	2	1	1	0	0
	Post-	3.47	1.05	0.147	6	3	3	0	2	2	1	2	0
LOMPOC-IV	Pre- 16	2.06			10	3	0	1	1	0	0	1	0
	Post-	4.13	2.06	0.018	3	1	3	2	3	2	0	1	1
LOMPOC-V	Pre- 25	2.28			13	3	4	2	2	0	0	1	0
	Post-	3.92	1.64	0.006	11	0	3	0	1	2	4	3	1
LOMPOC-VI	Pre- 21	2.19			15	0	1	2	1	0	1	1	0
	Post-	3.33	1.14	0.148	10	1	2	3	0	1	0	2	2
TOTAL	Pre- 142	2.58			76	19	11	7	8	6	7	5	3
	Post-	3.70	1.12	0.000	56	9	13	10	11	12	9	15	7

Table 13

EXPERIMENTAL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT--INDIVIDUAL STANINE SCORES-- SENTENCE STRUCTURE

SCHOOL	SAMPLE SIZE	MEAN SCORE	STANINE GAIN	z-RATIO	Below Average				Average		Above Average		
					1	2	3	4	5	6	7	8	9
LOMPOC-I	Pre- 18	5.50			3	0	4	0	0	0	6	3	2
	Post-	6.56	1.06	0.219	0	1	2	0	3	1	2	6	3
LOMPOC-II	Pre- 43	4.63			3	9	8	3	5	2	2	8	3
	Post-	6.91	2.28	0.012	1	1	1	2	4	4	9	13	8
LOMPOC-III	Pre- 19	4.79			2	2	4	0	2	4	2	2	1
	Post-	6.89	2.11	0.024	0	0	1	0	3	2	4	8	1
LOMPOC-IV	Pre- 16	3.81			3	4	2	2	0	2	1	1	1
	Post-	7.38	3.56	0.034	0	1	0	1	0	1	2	7	4
LOMPOC-V	Pre- 25	3.84			8	2	4	3	0	2	1	4	1
	Post-	6.64	2.80	0.021	1	2	1	0	2	4	2	7	6
LOMPOC-VI	Pre- 21	3.38			8	3	3	1	1	0	2	2	1
	Post-	5.86	2.48	0.020	2	0	1	3	4	2	2	3	4
TOTAL	Pre- 142	4.35			27	20	25	9	8	10	14	20	9
	Post-	6.71	2.37	0.000	4	5	6	6	16	14	21	44	26

Table 14

EXPERIMENTAL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT--INDIVIDUAL STANINE SCORES-- COMPOSITE (FINAL)

SCHOOL	SAMPLE SIZE	MEAN SCORE	STANINE GAIN	z-RATIO	Below Average 1	2	3	Average 4	5	6	Above Average 7	8	9
LOMPOC-I	Pre- 18	4.17			1	3	2	4	4	2	2	0	0
	Post-	6.28	2.11	0.029	0	1	0	1	4	4	4	2	2
LOMPOC-II	Pre- 43	3.74			2	12	9	6	6	3	4	1	0
	Post-	6.09	2.35	0.012	1	2	1	3	4	12	11	7	2
LOMPOC-III	Pre- 19	3.95			1	5	2	2	5	3	1	0	0
	Post-	5.53	1.58	0.022	0	0	1	4	6	2	4	2	0
LOMPOC-IV	Pre- 16	2.94			2	7	2	3	0	1	1	0	0
	Post-	5.63	2.69	0.034	0	1	1	1	3	6	2	2	0
LOMPOC-V	Pre- 25	3.20			3	8	3	4	6	1	0	0	0
	Post-	5.40	2.20	0.021	1	2	1	6	3	3	5	1	3
LOMPOC-VI	Pre- 21	3.19			1	10	4	0	3	2	1	0	0
	Post-	4.62	1.43	0.002	2	1	3	5	3	2	3	2	0
TOTAL	Pre- 142	3.56			10	45	22	19	24	12	9	1	0
	Post-	5.65	2.09	0.000	4	7	7	20	23	29	29	16	7

Table 15

CONTROL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT--INDIVIDUAL STANINE SCORES-- CONTENT

SCHOOL	SAMPLE SIZE	MEAN SCORE	STANINE GAIN	z-RATIO	Below Average			Average			Above Average		
					1	2	3	4	5	6	7	8	9
	Pre-	5.64			0	1	0	6	14	13	6	4	0
GROUP I	44												
	Post-	5.23	-0.41	0.147	0	1	1	13	11	12	3	2	1
	Pre-	6.06			0	0	1	7	6	6	6	5	3
GROUP II	34												
	Post-	5.32	-0.74	0.057	0	0	1	12	9	5	3	2	2
	Pre-	5.82			0	1	1	13	20	19	12	9	3
TOTAL	78												
	Post-	5.27	-0.55	0.020	0	1	2	25	20	17	6	4	3

Table 16

CONTROL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT--INDIVIDUAL STANINE SCORES--ORGANIZATION

SCHOOL	SAMPLE SIZE	MEAN SCORE	STANINE GAIN	z-RATIO	Below Average			Average			Above Average		
					1	2	3	4	5	6	7	8	9
GROUP I	Pre- 44	5.34			0	2	3	8	10	11	6	4	0
	Post-	4.64	-0.70	0.015	0	0	9	13	14	3	3	2	0
GROUP II	Pre- 34	5.62			0	0	6	3	11	3	3	5	3
	Post-	5.32	-0.29	0.522	0	1	1	14	6	3	3	3	3
TOTAL	Pre- 78	5.46			0	2	9	11	21	14	9	9	3
	Post-	4.94	-0.53	0.050	0	1	10	27	20	6	6	5	3

Table 17

CONTROL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT--INDIVIDUAL STANINE SCORES--MECHANICS

SCHOOL	SAMPLE SIZE	MEAN SCORE	STANINE GAIN	z-RATIO	Below Average 1	2	3	4	Average 5	6	Above Average 7	8	9
GROUP I	Pre- 44	4.93			9	2	4	5	4	4	5	6	5
	Post-	4.43	-0.50	0.401	10	6	4	1	5	4	5	8	1
GROUP II	Pre- 34	5.03			5	3	3	2	4	5	6	3	3
	Post-	3.88	-1.15	0.069	7	9	4	1	2	2	4	4	1
TOTAL	Pre- 78	4.97			14	5	7	7	8	9	11	9	8
	Post-	4.19	-0.78	0.074	17	15	8	2	7	6	9	12	2

Table 18

CONTROL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT--INDIVIDUAL STANINE SCORES-- SENTENCE STRUCTURE

SCHOOL	SAMPLE SIZE	MEAN SCORE	STANINE GAIN	z-RATIO	Below Average			Average			Above Average		
					1	2	3	4	5	6	7	8	9
GROUP I	Pre- 44	6.09			6	1	0	1	4	9	8	9	6
	Post-	6.36	0.27	0.567	0	1	3	5	5	5	12	7	6
GROUP II	Pre- 34	7.29			1	0	1	1	0	4	7	13	7
	Post-	6.09	-1.21	0.006	0	2	0	5	7	5	4	9	2
TOTAL	Pre- 78	6.62			7	1	1	2	4	13	15	22	13
	Post-	6.24	-0.37	0.271	0	3	3	10	12	10	16	16	8

Table 19

CONTROL GROUP RESULTS OF SANTA BARBARA COUNTY EXPERIMENT--INDIVIDUAL STANINE SCORES-- COMPOSITE (FINAL)

SCHOOL	SAMPLE SIZE	MEAN SCORE	STANINE GAIN	z-RATIO	Below Average			Average			Above Average		
					1	2	3	4	5	6	7	8	9
	Pre-	5.30			1	3	2	7	8	13	6	4	0
GROUP I	44												
	Post-	5.02	-0.27	0.416	0	1	5	12	8	11	6	1	0
	Pre-	5.94			0	1	1	6	4	8	9	3	2
GROUP II	34												
	Post-	5.15	-0.79	0.043	0	0	5	12	2	7	5	2	1
	Pre-	5.58			1	4	3	13	12	21	15	7	2
TOTAL	78												
	Post-	5.08	-0.50	0.056	0	1	10	24	10	18	11	3	1

Table 20

PRE-POST TEST ACHIEVEMENT CHART: COMPOSITE (FINAL) GROUP SCORES

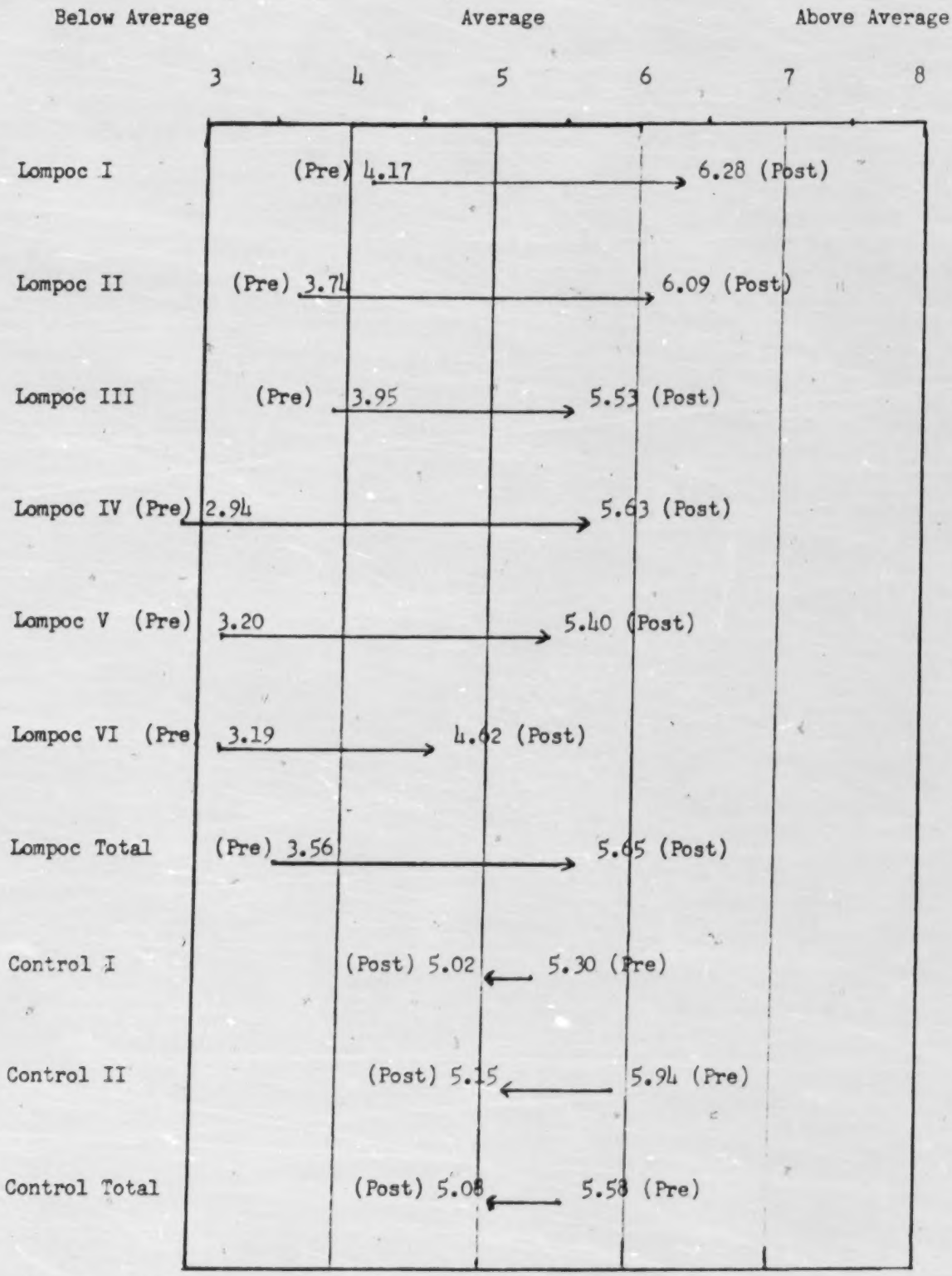

Table 21

VII. DISCUSSION OF RESULTS

Tables 1 through 5 list the group results in stanine scores of the experimental classes for all items tested: content, organization, mechanics, and sentence structure as well as a composite or average (final) score of all items tested. Tables 6 through 10 list the group results for the control classes for all items tested. Tables 11 through 15 list the individual stanine scores of the experimental classes for all items tested. Tables 16 through 20 list the individual stanine scores of the control classes for all items tested. Table 21 contains a chart portraying the composite or final achievement in writing skills of all groups tested.

Tables 1-5

Tables 1 through 5 list not only the group results in stanine scores of the experimental classes for all items tested--content, organization, mechanics, and sentence structure--but also the sample size, maximum score, minimum score, range, mean, variance, standard deviation, mean deviation, median, mode, stanine gain, z-ratio, and significance or confidence level of the results. Classes are arranged by period or time block.

Table 1

Table 1, dealing with experimental group scores for content, reveals a highly significant gain for all classes except VI. The stanine gain ranged from 0.71 for class VI to 2.93 for class II. The mean gain for the six groups(classes) as a whole was 2.28 at a significance level of 0.000, thus indicating a 100 percent probability that the gain was attributable not to random chance but to treatment of the groups as a whole.

One point should be made to clarify the meaning of the z-ratio and the significance of the results. The z-ratio reveals whether the gains--the difference in the means between the pre-and-post-test scores--are significant--that is, attributable to treatment (teaching) or to random chance. The smaller the number of the z-ratio, the less likelihood there is for random chance, and, therefore, the more significant are the results, because the higher is the probability that the results are attributable to teaching. For example, a z-ratio of 0.01 (significant at the 1% level of confidence) means that there is only a 1% probability of the increase being attributable to random chance. This is more significant than a z-ratio of 0.05, where there is a 5% probability of the increase being attributable to random chance, and much more significant than a z-ratio of 0.10 (10%), which is really not considered to be significant by most researchers. In short, results are considered most significant at 0%, very highly significant at 1%, very significant at 5%, and least significant at 10%. Above 10% they are not considered significant at all.

It might be noted that class II, containing the largest sample--and the only one in the experimental group to include an entire class--revealed the greatest gain of all. That gain, however, was almost identical to the gain revealed by class I, a class containing one of the smallest samples tested. Since both classes had the same instructors who used basically the identical methods, one might infer, first of all, that the "representative" sample was reliable, and secondly that these instructors placed a considerable emphasis on content, perhaps more so than their peers.

Comparing the pretest mean scores in content for all groups, one notes a clustering of scores among four classes--I, II, III, and VI--and another

cluster between two other classes--V and VI. One might therefore infer that the population samples were fairly homogeneous among these groups at the beginning of the course even though there had been no attempt to group the students by ability. (As explained earlier, however, under METHOD, the school's lowest ability students, numbering seven percent, were not included in the regular freshman classes that were tested.) Examining the dispersion of scores, one notes a difference of only 1.45 stanines between the highest and lowest scoring groups on the pretest, and a greater dispersion (2.57) on the post-test, thus indicating that achievement may vary among groups regardless of their initial similarities.

Table 2

Table 2, dealing with organization, reveals highly significant gains for all classes (in the experimental group), gains ranging from 1.57 stanines for class VI to 3.56 for class I. The significance level was 5% for each of the classes and 0% for the group as a whole. Again, for the group as a whole, this gain must be attributable to treatment by a probability of 100%. Concerning the clustering of scores, it might be noted that all classes except one (III), scored below average on the pre-test; on the post-test, however, there was more of a spread, with two classes scoring in the average range, three in the high average, and two above average. Whereas the spread between the highest and lowest class was only 1.01 on the pretest, it had increased to 2.23 on the post-test, again indicating variations in achievement among groups. The most obvious factor in table 2 is that all of the experimental classes achieved, thus indicating that every instructor emphasized organization and practically every student learned it.

Table 3

Table 3, dealing with mechanics (spelling and punctuation), reveals that only two classes--IV and VI--of the six in the experimental group achieved significant gains in this item. Class I and II showed almost no gain at all; class III and VI showed a gain of over 1 stanine, but the z-ratio indicated a high probability that the increase may have been attributable to random chance; and class IV and V revealed a gain of 2.06 and 1.64 stanines respectively, at a significance level of 5% and 1% respectively. For the groups of a whole, however, because of their large sample size, a gain of 1.12 stanines was revealed at a significance level of 0%, thus indicating that highly significant gains were achieved.

Concerning the dispersion of scores, one notes that all groups scored below average on the pretest, with a spread of 1.55 stanines separating the highest scoring group from the lowest. On the post-test, the spread was 1.73, not significantly different at all. Furthermore, at the conclusion of the course, only three classes--I, IV, and V--had moved above the below-average range, and even then, they were barely average, since none had scored above 4.13.

Despite the significant gains made by the two classes mentioned, the tables indicate that the students lacked spelling and punctuation skills upon entering their classes, and in effect still lacked them at the conclusion. For even though the group as a whole achieved 1.12 stanines at a 100% probability, the post-test score of 3.70 indicates that the group was still below average in this area.

Table 4

Table 4 deals with sentence structure, which includes not only correct grammar, usage, and syntax, but also appropriate diction and variety in

sentence patterns. Only one class--I--did not show a significant gain in this area, even though it did achieve a gain of 1.06. All other classes revealed highly significant gains (at 5%) ranging from 2.11 to 3.56 stanines. The mean gain for the experimental group as a whole was 2.37 stanines at 0%, thus indicating a highly significant gain in sentence structure. Moreover, of all the items tested, both on the pre-and-post-test, sentence structure revealed the highest scores among the experimental classes. Pretest scores ranged from 3.38 to 5.50, with three classes in the below average range and four in the average range of 4.35 or better. Post-test scores ranged from 5.86 to 7.38, with only one class in the average range, four in the high average or very high average, and one above average. The post-test spread of 1.52 between the highest scoring class and the lowest was smaller than that on the pretest (2.12), thus indicating less individual differences among classes in this item.

Table 5

Table 5 deals with the composite or final scores of the experimental classes. This "final" score is an average of the four other items already discussed. It should be noted that all classes achieved highly significant gains in their final scores, ranging from 1.43 stanines at a significance level of 1%, to 2.69 stanines at a significance level of 5%. The mean achievement of the group as a whole was 2.09 stanines at 0%, thus indicating that all of the classes tested in the experimental group did, indeed, improve in their writing skills.

The pretest reveals a cluster of scores, with five classes in the below average range and one in the low average. The spread between the highest scoring group and the lowest was 1.23 stanines. At the conclusion of the course, all classes but one revealed at least average scores, with

the lowest scoring class in the low average range, three in the average range, and two in the high average range. The spread between the highest scoring class and the lowest was 1.66 stanines, not significantly greater than that revealed eighteen weeks earlier by the same group, thus indicating more of a similarity among groups in achievement than individual differences.

Tables 6-10

Tables 6 through 10 list the group results in stanine scores of the control classes for all items tested, as well as the composite or final score. Tables 6, 7, 8, and 9 deal with content, organization, mechanics, and sentence structure respectively. Table 10 deals with the final composite score. Upon examining these tables, one notices that in each item tested the control group as a whole reveals a loss rather than a gain. Although the loss is so slight as to appear almost negligible, the significance of the results, especially in tables 6 and 7 (content and organization), indicates that whatever loss there is cannot be attributable to random chance but rather to treatment of the classes as a whole. Table 8, revealing a loss in mechanics, indicates a high probability of random chance: 10%. Table 9, dealing with sentence structure, indicates that although one class (III) revealed a significant loss, the two classes as a whole may or may not have lost their skills; the results were not significant. Finally, table 10, dealing with the final score, reveals similar findings to those in the preceding table; that is, although one class (II) showed a loss, the overall loss of the groups as a whole was not significant. One class (II) revealed a loss in four of the items tested: content, mechanics, sentence structure, and the composite (final) score. However, the only truly significant loss for that group appears

to be in sentence structure, where the loss is actually more than one stanine at a significance level of 1%.

Tables 11-15

These tables list the individual stanine scores of the experimental classes in all items tested, as well as the mean scores for each class, the stanine gain, and the z-ratio. Although a detailed discussion will not be presented, the researcher would like to make some observations. If a curve were to be drawn to represent the individual scores on the pretest for all items tested, the curve would be skewed sharply to the left. In content, organization, mechanics, sentence structure, and the final score, the percentage of below average scores would be as follows: 45%, 50%, 74%, 50%, and 54%. The percentage of above average scores on the pretest for these items would be 7%, 6%, 10%, 30%, and 7%. Conversely, the post-test scores indicate a curve skewed sharply to the right except for mechanics and the final score, as follows: 50%, 55%, 21%(mechanics), 64%, and 16%. However, even in mechanics and in the final score, the curve is skewed to the right, but not so sharply. A definite achievement in all areas is therefore evident.

If one examines the curve depicted by the pretest composite(final) scores, he notices a broad skew to the left, as follows: 54% below average, 39% average, and 7% above average. This skew reveals below average ability in overall writing skills. At the conclusion of the course, however, the post-test scores reveal achievement in overall writing skills, because the curve is now bell-shaped, with 13% in the below average range, 50% average, and 17% above average. The greatest shift has been from the below average range to the right, or upwards. Forty-one percent of the students who had exhibited below average writing skills on the pretest have improved signifi-

cantly, some of them even dramatically, for they exhibit at least average skills, or in some cases skills that must be considered above average. In short, the overall achievement in writing skills of the experimental group is obvious.

Tables 16-20

These tables list the individual scores of the control classes in all items tested, as well as the mean scores for each class, the stanine gain, and the z-ratio. A detailed discussion will not be presented, but the researcher would like to make the following observation, concerning the composite (final) scores. Pretest scores plotted on a curve would reveal a skew slightly to the right, with 10% below average scores, 58% average, and 30% above average. The curve represented by the post-test scores would not have changed appreciably, except for a greater number of scores in the average range (66%) and fewer in the above average. Very little if any achievement is therefore evident.

VIII. CONCLUSION

The statistics confirmed all of the hypotheses except one, which concerned achievement in all classes--both experimental and control. The experimental classes did achieve, but the control did not.

First a significant range of abilities was evident among students within each group or class, as well as among the various classes, both in the pre-and-post tests. This range reflected the various writing skills tested, as actually demonstrated by the students in their writing. In content, for example, the range of the experimental students varied from 5 stanines on the pretest to 8 on the post-test, as revealed in table 1. For the control group (table 6), this range in content varied from 6 to 7.

Of all the skills tested, including the overall writing skill referred to as the composite (final) score, the smallest range evident among the experimental classes was revealed in content, organization, and in the final score: 5. The largest range (8) was evident in every item. For the control classes the smallest range (5) was revealed in organization; the largest (8) in mechanics, in sentence structure, and in the final score.

The second hypothesis was partially confirmed because the post-test did in fact reveal a significant range of achievement in most areas tested--but only for the experimental classes. The control classes showed no achievement at all in any of the areas tested. The achievement for the experimental classes as a whole was most significant in each of the areas tested. This group achievement ranged from 1.12 stanines in mechanics at a significance level of 0% to 2.61 stanines in organization, also at a significance level of 0%.

The third hypothesis was fully confirmed. Because of the Kates Method of individualized instruction, the experimental classes achieved greater and more significant gains than did in the controls, in all areas tested. The control classes, using the traditional method of instruction, not only failed to show any achievement in any of the areas tested, but in fact even showed a loss in some areas. The achievement of the experimental group was attributable to treatment of the classes by a probability of 100%. The treatment used was the Kates Method of instruction.

The final hypothesis was confirmed, thus supporting the researcher's previous study. The overall mean achievement of the experimental classes, based on the composite (final) score, was comparable to that exhibited by the entering college freshmen who had been tested by the researcher three

years earlier, during the Kates Survey of College Freshman Composition Writing Skills: Sixteen First-Semester English Composition Classes from Eleven Community Colleges and Universities, ERIC, ED 077 498, June 1973. The mean achievement of the 142 entering ninth graders at Lompoc Senior High School was 2.09 stanines at a significance level of 0%. The overall mean achievement of the 285 entering college freshmen enrolled in 16 different classes among 11 colleges and universities was 1.804 stanines, also at a significance level of 0%.

Because the final hypothesis was so fully confirmed, even though the population sample of ninth graders used in the experiment could hardly be compared to that used in the researcher's previous study dealing with college freshmen, certain conclusions must now be drawn. Although these conclusions will be drawn from this particular study, comparisons will be made occasionally to the previous one.

First, in any given population sample, but especially in a heterogeneous sample--regardless of similarities in age, socio-economic background, or previous education--there will exist considerable or even vast differences in writing abilities, as demonstrated by the students on an essay test. These abilities or skills can be measured with an extremely high degree of accuracy by anyone trained in using the Kates Grading Scale. The reader so trained can reveal data, which can later be computed to determine--again with a high degree of accuracy--the actual achievement in writing skills, not only of the individual student but also the achievement of groups of students and entire population samples. In short, individual differences in writing skills do exist, they can be analyzed, and they do account for differences in achievement.

Second, a very reliable population sample for experimental purposes can be drawn, by random chance, from a large population sample. The experimental sample need be no larger than one-third of the original. To illustrate, the results of the experimental class II were almost identical to those of class I, taught by the same instructors, even though the sample size of class I was only 41% of the larger sample size.

Third, a correlation does exist between the instructor and the treatment of the class. Even though a group of instructors may use the same basic method of instruction, individual differences exist among instructors in their attitude toward the same subject, and in their emphasis on different items, all of which often results in varying achievement among classes. To illustrate, all six instructors in the experimental classes used the same basic technique: the Kates Method; all assigned the same number of themes requiring the same type of expository organization; yet only two classes (IV and V) achieved significant results in mechanics. Both of these classes were team taught, as were all the others. However, the variable factor appears to be one instructor who was common to both of these classes. He used a different approach to the conference method during both classes, as did the other instructor involved. Since his classes were the only ones to receive conferences both on the first and final drafts rather than on the final one alone, and since his classes were the only ones to achieve significant results in mechanics, the achievement appears to be attributable to his particular treatment.

Fourth, the converse is also true: despite individual differences, instructors can use the same basic technique and achieve comparable results. The variable, of course, is the population sample. However, even that variable will not necessarily make the difference. The Kates Method,

explained earlier, was used by all instructors in the experimental classes, yet despite the instructors' individual differences and even the differing population samples among classes (they were not homogeneous), all of the experimental classes achieved highly significant or most highly significant gains in their composite (final) score. Furthermore, the one particular item that was emphasized most by the department head--organization--was the only item in which every class achieved significant gains. The common denominator was the Kates Method, with a special emphasis on one item: organization. In short, even though individual differences among instructors may account for a different emphasis in teaching one or more aspects of the same subject, the instructors can use the same basic technique, they can emphasize the same item, and they can expect to achieve comparable, but not identical, results. In other words, instruction can be standardized to a very great extent.

Fifth, since the treatment of any class is largely dependent upon the conditioning of its students by the instructor, it appears that the "tone" or atmosphere of the class is largely set by the individual instructor at the very beginning of the course, and this atmosphere will affect achievement. This correlation was emphasized by all the instructors of the experimental classes. They promised their students success if the students cooperated and worked hard, and throughout the course they continued to maintain a "serious" attitude toward the subject, asking their students in turn to take the subject "seriously." The initial tone was undoubtedly reinforced at the beginning of the course by the consultant--an "outsider"--who not only addressed each class but also provided most of the students with their first "individual" conference. Writing, therefore, was a serious matter, and all writing--graded or not--was to be taken seriously. The tone, initiated and maintained by the instructors, appeared to contribute toward

the motivation of the students. The conferences, besides providing meaningful feedback, also provided encouragement, and thus additional motivation. Finally, the conferences perpetuated the tone: writing was important. In short, the tone became part of the conditioning treatment.

By contrast, the instructors in the control classes admitted that the atmosphere in their classroom may at times have been almost too relaxed. Their students, when given the post-test, asked whether they would receive a grade for it, and when they learned that they would not be given credit, they appeared to consider the test merely another ungraded assignment. This "attitude" may have been one of the variables in the results. Motivation, obviously, was a factor.

Fourth, the achievement of a particular class cannot be determined readily until after the pre-and-post testing has been completed. Classroom control problems are not necessarily an indication of lack of achievement. Commenting on a particular experimental class, the instructor felt that "it was not a good class; there were too many discipline problems." Despite that assessment, however, the class (II) achieved highly significant gains in every item except in mechanics. Moreover, the gain in the composite (final) score was the second highest of all classes tested. On the other hand, lack of classroom control problems are not an indication of achievement, either. In the control classes, the atmosphere, as already mentioned, was not only relaxed but warm and personal, according to the instructors. The students were given an adequate number of writing assignments, which they completed. Yet none of the classes showed any achievement whatsoever. The conclusion must be that despite the lack of outwardly disruptive behavior, they either failed to learn the skills that were tested, or they were not motivated to perform on the post-test. It appears, therefore, that classroom control problems, or lack of them, are not the

key elements in the successful treatment of a class.

The key element, according to all of the instructors--both experimental and control--was the individual conference. Regardless of the "rapport" between the instructor and student, which was evident in the control classes, one vital element was missing: meaningful feedback, available only through the two-way oral interchange of the individual conference. (The returned, graded paper containing symbols and written comments by the instructor may or may not communicate his meaning; invariably it does not, for the student often continues to commit the same errors, and in fact he often crumples up the paper and throws it away in disgust. He has been intimidated or even humiliated by the red marks and negative criticism. This widely observed phenomenon was reported in the researcher's previous studies.) Students in the control classes did not receive personal conferences, and like other control students in previous studies conducted by the researcher, they did not achieve as significant gains in their writing as did the experimental students, who did receive conferences. The students in the experimental classes were free to question their instructors during the conferences, but regardless of whether they took advantage of that opportunity, they did receive meaningful feedback--and encouragement. The human element was always present in the feedback. As one of the experimental instructors emphasized, "Conferencing was the key. Even the low achievers seemed to improve drafts after conferencing."

In summary, this study concludes that the student's improvement in writing skills appears to be correlated positively to three factors: one, the amount of expository writing done in class; two, the amount of individual conferencing the student receives as feedback on his writing; and three, the motivation provided by the instructor, probably as a result of

his initial attitude toward the teaching of the subject--an attitude that must prevail throughout the course. The first two of these conclusions are identical with those drawn by the researcher in his earlier study of college freshman writing skills during the Kates Survey (ERIC, ED 077 498, June 1973). It appears, therefore, that they are valid. The third one is new and should invite further research.

One conclusion drawn from the earlier study appears open to question, or at least open to revision. It concerns the correlation between the "intimacy of the class itself, resulting from a smaller class size" and achievement. This study would appear to contradict that earlier conclusion because all of the classes tested were fairly large, ranging from 38 to 48, and the "intimacy" factor, therefore, appeared to be absent. Despite the large size, however, it must be emphasized that because the classes were all team taught, the student-teacher ratio was never more than 24:1, and actually as low as 19:1, virtually the same ratio recommended in the previous study. Furthermore, in this experiment, at least one teaching assistant was usually present in each experimental class, to lighten the burden of classroom control and to take care of minor, time-consuming details, thus relieving at least one instructor so that he could provide individual conferences, often without interruption.

The final conclusion of this study is identical to that of the previous one dealing with college freshmen, except for the underscored qualification which has been added. Ultimately, almost every student is capable of achieving--under certain conditions--and his achievement, as well as his probability for success in the course, can be measured with a very high degree of accuracy by the utilization of essay tests, grading procedures, and computations similar to those utilized in this study.

IX. SUMMARY

In order to test the reliability of the Kates Method of individualized instruction (in the teaching of English composition) on high school students, and in order to support or challenge previous work on individualization, a pre-and-post essay test was administered to six classes of experimental ninth-graders, numbering 142, at Lompoc Senior High School, and to two classes of control ninth graders, numbering 78, at a neighboring school, both located in Santa Barbara County, California. Rigidly controlled conditions prevailed throughout both tests and throughout the grading, which was done by an "outside" reader, especially trained for the project. Grading criteria on the Kates Grading Scale included content, organization, mechanics, and sentence structure, all of which were evaluated equally on the stanine. The composite (final) score was an average of the other four items.

Throughout the semester's course, each experimental class was assigned approximately fourteen essays, varying in length from one to five paragraphs, mostly thesis-and-support in nature. All were written in class. Each student received a five-minute conference on alternate weeks. The control group, less structured, received ten to twelve paragraph writing assignments, to be done in-and-out of class, but received grades only on four or five. Feedback was by written comments, no conferences.

The results, utilizing the 2 tail test and z-ratio, confirmed all the hypotheses except the second, which was partially confirmed. First, a significant range of writing abilities was evident within each class, thus reflecting the various skills tested. Second, the post-test revealed significant gains in most areas tested--but only for the experimental classes. The control showed no gain at all. The mean group achievement

of the experimental classes ranged from 1.12 stanines in mechanics to 2.61 in organization, both at a significance level of 0%, thus indicating no possibility of random chance. Third, because of the Kates Method of instruction, requiring regular, sequential writing assignments and regular teacher-student conferences, the experimental classes achieved greater and more significant gains than those of the control, who did not achieve at all. The experimental group achievement in the composite (final) score was 2.09 stanines at a significance level of 0%, very comparable to the overall achievement of 285 entering college freshmen three years earlier (1.80 stanines at 0%), in the Kates Survey of College Freshman Composition Writing Skills (ERIC, ED 077 498, June 1973). The achievement of the experimental students was attributable to treatment of the class by a probability of 100%, thus confirming the superiority of the method.

X. RECOMMENDATIONS

After having analyzed his study, the researcher would like to make the following recommendations, which he thinks should apply particularly to the teaching of English composition at the junior and senior high school levels.

First, all English classes should utilize pre-and-post essay tests, not only to diagnose the individual's ills and skills at the very beginning of the course, but also to determine the achievement in writing skills of each individual and of the class as a whole. If the Kates Essay Test and Grading Scale are utilized, the statistics are likely to prove reliable, especially if the reader or readers are trained in the process. Furthermore, if one is able to utilize the services of a computer-programmer, the statistics can then be used for educational research.

Second, because the "tone" or atmosphere of the class often has a direct effect upon the achievement of all students, and because this tone is often determined largely by the instructor's "initial" attitude toward his subject--the teaching of specific skills--the instructor must be very careful to make the appropriate impression upon his students during the first class meeting or at least during the first week of class.

Third, a logical sequence of instruction in the teaching of sentences, paragraphs, and larger themes ought to be utilized for maximum effect. There is little point in attempting to teach a complex organizational pattern, like the comparison/contrast essay for example, if the student has not yet mastered the simpler forms of organization, like process analysis or thesis-and-support. Nor is there any point in teaching complex rules of grammar if the student has not yet mastered the simple sentence.

Fourth, regular, weekly writing assignments of paragraph length or more should be manditory for all students, and as much writing as possible should be done in class, not only to ensure the authorship of the writing, but also to provide a reasonably quiet atmosphere that will enable the instructor to provide his students with individual conferences on their written work.

Fifth, each instructor should provide for every student in class a regular series of conferences on the written assignments, each conference lasting approximately five minutes. Because the logistics of large classes and short time blocks usually preclude such a conference every week, the instructor should arrange his teaching schedule so that each student receives a conference every other week. During the conference, the instructor should furnish not only feedback on the written work but also encouragement for the student. This particular recommendation is virtually identical to the one made by the researcher three years earlier in his previous study on college freshmen composition writing skills. This recommendation, more than any other, should be manditory, because the evidence supporting it is conclusive.

Sixth, students who are particularly weak in mechanics, or in any other area, ought to receive additional instruction in that area. The instruction might be in the form of individualized conferences, but individualized exercises for reinforcement should also be incorporated.

Finally, a continuing, logical sequence of expository writing assignments, in conjunction with regular teacher-student conferences, ought to be manditory for all students at all grade levels in all schools. This recommendation is vitally important because the process of learning writing skills is unusually complex and often extremely difficult to master; and if the skill is not practiced, it is easily forgotten.